

# International Life Skills Survey

## Team Leaders Meeting

August 23-25, 1999

Princeton, NJ

## SUMMARY NOTES

### *Introduction and Overview*

Scott Murray and Marilyn Binkley gave an introduction, outlining the goals of the meeting, and detailing the current state of affairs surrounding ILSS. The need to make critical decisions over the next few weeks was outlined, as between this meeting and the NPM Meeting in September the readiness of each domain for inclusion in the pilot test must be decided upon. After the NPM Meeting, there will be country-level commitments to create national planning reports, review the frameworks nationally, and contribute to the production of new items. The time frame for new item development, building constituencies around the frameworks, overlap among the domains, and quality assurance were all topics that were identified as open for discussion.

### *Review of Procedures*

There was a discussion on the criteria for deciding what items will go into the pilot survey. It was decided that these criteria would be presented at the NPM Meeting in the form of a memo to be constructed by ETS. The criteria discussion included the finality of the frameworks presented at the NPM Meeting, item development and mapping, purpose of the feasibility study, preparations for the pilot survey, and country-level input into final item pools. The option of utilizing cognitive labs/focus groups for further study of items was considered as well.

### *Teamwork*

David Baker gave a presentation on the Teamwork domain, highlighting the framework, item development, initial feasibility study results, and next steps. Questions were raised regarding the cultural specificity of the domain and the literacy level required for the assessment. There was an extensive discussion of scoring and scaling this module as well, focusing on the possibility of right/wrong scoring along with possible means of scaling. The importance of completing the feasibility study in a second country was noted.

### *Literacy*

Irwin Kirsch gave a presentation on the Literacy domain. Two main activities were discussed, revising the framework and the development of a new item set to replace those from IALS. There was a discussion of the differences between Prose and Document

Literacy and the nature of the two domains. The development of new items through multiple language networks was detailed as a means for acquiring new items and expanding the domain constituency. The issue of copyright protection was addressed as it relates to literacy items. The item development and translation processes employed by the literacy domain were considered as a possible model for the other domains, but the unique nature of the literacy domain as an established assessment was noted to preclude any simple comparisons.

Since some of the item development teams are behind schedule, it was noted that a decision will have to be made as to the final timeframe for new items, and what level of inclusion and influence over new items will be allowed after the final due date.

### *Numeracy*

Iddo Gal gave a presentation on the Numeracy domain, focusing on the transition from Quantitative Literacy to Numeracy, the current framework, item development, and the feasibility study results. There was a discussion of issues surrounding overlap between this scale and the Document Literacy scale, item difficulty, scoring methodology, building constituencies, and developing new items. It was noted that for political reasons, it will be necessary to allow for country-level input into new item development.

### *Overlap of Instruments*

Marilyn Binkley led a discussion on the issue of overlap among the domains, as the issue of overlap is not specific to the Numeracy and Document Literacy scales, but applies across all domains to some degree. Options such as using individual items on multiple domains, individual items on one domain, and various scaling, scoring, and reporting possibilities were all discussed extensively. It was noted that given variables never fully account for item difficulty, and overlap will exist in some sense no matter what when dealing with cognitive assessments. Participants decided that ultimately, overlap is a reporting issue, and this is where the focus should lie, in terms of both general and technical reporting.

### *Problem Solving*

Anne Hensgen and Judith Ebach gave a presentation on the Problem Solving domain, highlighting the framework, item development, and the feasibility study results. The length of the assessment tasks in relation to the nature of the tasks and how they are constructed was discussed extensively. Scaling and scoring methodologies for the domain were also discussed, particularly in terms of developing proficiency levels, conducting the feasibility study, and in relation to the literacy and numeracy domains.

The focus of the Problem Solving domain on analytical problem solving was considered in relation to conceptions of general problem solving ability. The results from the Swiss interview study will hopefully provide further illumination as to the nature of the domain, as will the results of the PISA problem solving assessments. It was noted that the name

of the domain should more accurately reflect the focus on analytical reasoning to circumvent potential problems in the future.

### *ICTL*

Jean-Paul Reeffer gave a presentation on the ICTL domain, concentrating on the current planning that is going into developing the domain. The shift away from a notion of computer familiarity towards a more expanded notion of communication was detailed. Constituency building prior to framework development was laid out as the preferred means of development for this domain, with assessment expertise only being brought in at the end. A more pragmatic than scientific approach will be taken.

Issues surrounding what exactly will be assessed, how to account for rapidly changing technology and contexts, the type of assessment to be given, and the type of results that can be expected were all discussed.

### *Quality Assurance*

Stan Jones led a discussion on quality assurance, focusing on data collection, process control, sampling/weighting, item adaptation and psychometrics, and scoring. The background of the ILSS Quality Assurance team was described, and its two tasks given as providing recommended guidelines to the PMT, and following the acceptance of some form of those guidelines, monitoring ILSS processes and procedures. The developed guidelines are heavily based on team experience with IALS, and are meant to eliminate and minimize potential problems.

Criteria for the feasibility and pilot studies are being developed to apply generally to all scales. The importance of documentation and defining the critical item elements was stressed in relation to adaptation and translation. The work of Axel Buchner with the Problem Solving development team on translation and adaptation issues was described and discussed. The importance of the pilot study to further understanding these issues and for future decisions was noted.

Development team involvement in the entire translation and adaptation process was stressed throughout the discussion. The highly political nature of these processes was noted, as was the fact that perfection in these matters is unattainable. Cost implications were considered as well.

### *Developing a Consolidated Administrator's Manual*

Jean Pignal led a discussion on the process of developing the administrator's manual for ILSS. It was noted that this topic relates back to detailed knowledge regarding the critical item elements and how each item works. Development teams will need to decide how they want their assessments administered, and Statistics Canada will devise uniform procedures. Issues relating to the potential responsibilities of an interviewer/administrator and associated cost implications were also considered.

### *Linkages with Quantitative Literacy*

There was a short discussion on the possibility of linking the Numeracy domain to the Quantitative Literacy domain of IALS, but it was decided that it would be best not to do so as the potential for abuse and misinterpretation is high.

### *Incentives*

Stan Jones led a discussion on the use of incentives to increase response rates, which has been debated among the Quality Assurance team. It was noted that while response rates may be improved, the question of psychometric impact must be dealt with, including the loss of trend data and comparability. Political and cost implications regarding the use of incentives were mentioned as well.

### *Background Questionnaire*

Trevor Williams gave a presentation on the current state of affairs surrounding the background questionnaire, including the current form of the questionnaire, changes initiated since IALS, additions to the questionnaire, and future decisions. The overarching need to limit the number of new items on the questionnaire was noted, especially in relation to the additional items coming from working groups and development teams. The political implications of the questionnaire and the need to justify the final variables were discussed at length.

The relationship between the background questionnaire and the skill assessments was detailed, and the “story” to be told by ILSS was discussed. The option of having a country-specific final module on the questionnaire was discussed extensively. The psychometric impact and the political implications of such an addition were considered as well.

Development teams were asked to submit their proposed additions to the questionnaire to Trevor Williams by August 27<sup>th</sup>.

### *Building Consensus*

Marilyn Binkley led a discussion on building consensus and constituencies around the domains. It was recognized that this effort will vary depending on the domain, and that there is no one model.

Iddo Gal described the current activities that the Numeracy development team has undertaken. Options including peer-level review, presentations at conferences, and feedback from organizations were all discussed. The current timeframe of ILSS and the near-finalized nature of the frameworks was noted as it relates to obtaining future feedback.

Four levels of review were identified: review of definitions, facets, items, and the extent that the items map back to the facets (item difficulty). Results of the review process should be kept internal to the project, and will be released as part of the final technical report. The review process should be coordinated with the ILSS timeframe, such that the first two levels can be done in the present phase, with the last two occurring between the pilot and main phases.

It was noted that ultimately, support for ILSS will hinge on the results, as political considerations will outweigh all others after the data is produced. However, it was also noted that due to this, the best tactic is to ensure that ILSS is scientifically strong, and to document the processes and procedures involved in the survey as much as possible.

### *Next Steps*

Scott Murray and Marilyn Binkley concluded the meeting with a run-down of the next steps to be taken. Development teams are asked to supply the final version of their frameworks to John Konstant by September 3<sup>rd</sup>, and in lieu of a revised framework, a short memo detailing their current thinking. Each development team contact will discuss upcoming work for the development teams with the team leaders. Development teams are asked to submit a one-page paper detailing a plan for initiating a framework and item review process to Marilyn Binkley by September 30<sup>th</sup>. Also, prior to the pilot study, each team should submit a draft report on the determinants of difficulty for their domain, and after the pilot study, a report on what the data showed. This last piece will eventually go into the technical report for ILSS.